CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector
Neuron, Vol. 44, 219–222, October 14, 2004, Copyright ©2004 by Cell Press

Generation of Recognition Diversity in the Nervous System

Minireview

Dietmar Schmucker[1,*] and John G. Flanagan[2,*]
[1]Department of Neurobiology
Harvard Medical School and
Department of Cancer Biology
Dana-Farber Cancer Institute
[2]Department of Cell Biology and
Program in Neuroscience
Harvard Medical School
Boston, Massachusetts 02115

For decades, it has been suggested that complex neural wiring might be specified by extensive diversity in receptor isoforms. Dscam is a cell surface protein with 38,016 potential alternatively spliced isoforms in the fly nervous system. Remarkable binding studies now show that Dscam isoform diversity indeed results in an unprecedented level of recognition diversity, showing isoform-specific homophilic binding. In vivo studies have begun to suggest models for use of Dscam diversity in neuron-target recognition, axon fasciculation, and neuron self-recognition.

The nervous system, like the immune system, needs to generate extraordinary diversity and specificity in recognition. There are approximately 10^{12} neurons in the human brain, interconnected by a complex and precise network of roughly 10^{15} synapses. By comparison, the genetic blueprint that must specify this wiring seems tiny; according to current estimates, containing only a few tens of thousands of genes. Although the gap is somewhat closer in simpler organisms, even in *Drosophila* there are approximately 250,000 neurons and only about 15,000 genes.

How then is the relatively small amount of genetic information unfolded to specify such a complex wiring diagram? Several mechanisms appear to contribute (Figures 1A–1C), including combinatorial use of multiple guidance cues, cell-cell recognition labels in gradients, and refinement of connections based on correlated firing activity of neurons (reviewed in Zhang and Poo, 2001; Dickson, 2002).

Another class of mechanism, suggested since the 1970s, could be based on analogy with immune recognition (reviewed in Boulanger et al., 2001). In the immune system, generation-of-diversity mechanisms produce a vast array of antibody and T cell receptor protein isoforms by reshuffling or mutating the genetic information. Could isoform diversification mechanisms similarly generate large-scale recognition complexity in the nervous system?

Intriguing information over the last few years has suggested that some families of neural recognition molecules may indeed have large-scale isoform diversity. Neurexins potentially encode thousands of protein isoforms generated by alternative splicing and alternative promoter usage. Although it is not yet known if this results in a comparable scale of recognition diversity, at least some of the alternative isoforms do affect ligand binding (reviewed in Missler and Sudhof, 1998). Cadherin-related neural receptors (CNRs or protocadherins) are encoded in an arrangement that is intriguingly reminiscent of the antibody genes, with 52 "variable" exons in humans, each encoding an entire extracellular and transmembrane region, grouped in three tandem arrays upstream of a single set of "constant" exons encoding the cytoplasmic region (Wu and Maniatis, 1999). Perhaps the most impressive example is the *Drosophila Dscam* gene, which has the potential for alternative RNA splicing to generate 38,016 isoforms—more than twice the number of genes in the *Drosophila* genome (Figure 1D). It has remained, however, unproven whether some or all of the Dscam isoform diversity actually translates into recognition diversity.

Remarkable results recently published in *Cell* by Larry Zipursky's lab strongly support this idea, finding that every Dscam isoform tested had distinct in vitro recognition specificity (Wojtowicz et al., 2004). Moreover, each isoform showed homophilic binding (that is, binding to the same isoform) and little or no heterophilic binding (to other isoforms), a finding that remains essentially unexplained but has fascinating implications for Dscam structure, evolution, and function. Two additional studies published recently in *Neuron* from the labs of Larry Zipursky and Tzumin Lee explore the in vivo function of Dscam and the potential requirements for diversity (Zhan et al., 2004; Wang et al., 2004). Taken together, these results lead to novel models for the generation of diversity and its role in neural recognition.

Discovery of Dscam and Its Isoform Diversity

Drosophila Dscam was first isolated in a biochemical screen for tyrosine-phosphorylated proteins interacting with Dock, an intracellular adaptor protein homologous to mammalian Nck (Schmucker et al., 2000; Worby et al., 2001). Both Dscam and Dock are functionally required for axon guidance and targeting in the developing embryonic nervous system (Schmucker et al., 2000). *Drosophila* Dscam was named after the first member of this immunoglobulin (Ig) superfamily class of receptors, the human protein DSCAM (Down Syndrome Cell Adhesion Molecule), positionally cloned on chromosome 21, as a candidate disease gene for mental retardation associated with Down syndrome (Yamakawa et al., 1998).

Characterization of the gene structure of fly *Dscam* led to the finding of a remarkable molecular diversity (Schmucker et al., 2000). The *Dscam* gene contains three arrays of alternative exons (exons 4, 6, and 9) that are combined with 20 constant exons and two alternative transmembrane domains. The resulting protein isoforms all have the same domain architecture (Figure 1D). This unusual combinatorial usage of alternative exons provides for high variability in Ig-like domains 2, 3, and 7 and a maximal repertoire of 19,008 different extracellular Dscam domains. Extensive expression analysis suggests that most of these different isoforms are indeed

*Correspondence: dietmar_schmucker@dfci.harvard.edu (D.S.); flanagan@hms.harvard.edu (J.G.F.)

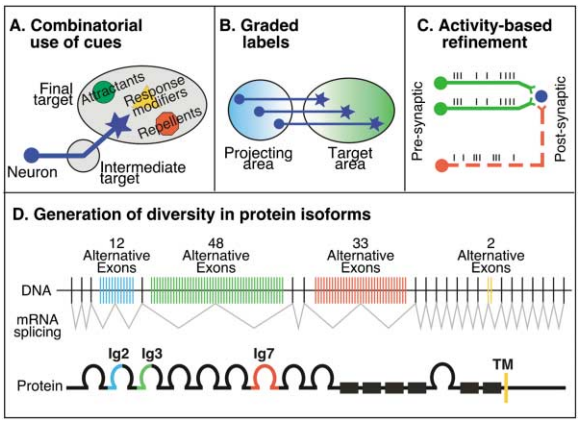

Figure 1. Mechanisms to Generate Complex Neural Connectivity from a Limited Number of Genes

(A) Combinatorially acting cues can be attractant or repellent guidance factors or can be modifiers that change how axons respond to these guidance factors. Combinatorial cues can act within the same location or sequentially: along pathways including intermediate targets, in final target recognition, and in synapse formation. (B) Graded labels provide an efficient way for a small number of cues to define a range of positional values across a large target area. Gradients can specify neural topographic maps, where axons find their position in the target based on position in the projecting area. (C) Neural activity can refine connections. Hebbian mechanisms are thought to reinforce connections if the pre- and postsynaptic neurons fire at similar times (green presynaptic neurons) or weaken connections if they fire at different times (red dotted neuron). (D) Some families of neural recognition molecules have the potential to generate large numbers of protein isoforms. The *Dscam* gene in *Drosophila* (upper line) has four arrays of alternative exons, with a single exon from each array being incorporated into each mRNA molecule. The resulting Dscam protein molecule (lower line) has ten Ig-like domains (semicircles), six fibronectin type III domains (boxes), and a transmembrane domain (TM). Alternative splicing creates diversity in the Ig2, Ig3, and Ig7 domains.

expressed in vivo (Schmucker et al., 2000; Hummel et al., 2003; Neves et al., 2004; Zhan et al., 2004). This extraordinary diversity inspired the theory that Dscam might represent a molecular code directing synaptic specificity (Schmucker et al., 2000).

Dscam Binding: Specific and Homophilic

It has, however, remained unproven whether all—or indeed any—of the isoform diversity of Dscam actually results in binding specificity. This question is now addressed in a landmark study by Wojtowicz et al. (2004).

Their study takes as a starting point observations on the mammalian homologs of Dscam, which do not show extensive alternative splicing and, like many other IgCAMs, bind homophilically (Agarwala et al., 2001). Wojtowicz et al. therefore tested for homophilic binding of fly Dscam. This experiment, and their other binding studies, all used soluble fusion proteins consisting of the Dscam ectodomain, fused to an immunoglobulin Fc tag and coated onto beads. These beads were then tested in three assays: self-aggregation, binding to cells expressing full-length Dscam, or coprecipitation of full-length Dscam from detergent-solubilized cells. These assays are all qualitative, and it would be informative in future to know the binding affinities. Nevertheless, the assays appeared to distinguish very cleanly between isoforms that bound and those that did not. The results of the initial experiments showed that two arbitrarily chosen isoforms did indeed bind to themselves homophilically.

Next came the crucial tests for specificity. These were done in an impressively systematic set of experiments involving 11 distinct isoforms in multiple pairwise combinations, including 5 of the 12 alternative Ig2 domains, 5 of the 48 Ig3 domains, and 7 of the 33 Ig7 domains.

The first specificity test was relatively nonstringent, involving two isoforms that differed extensively in all three variable domains. No heterophilic binding could be detected, demonstrating homophilic specificity for this pair of isoforms.

A more stringent set of specificity tests followed, using isoforms differing in only a single variable domain: that is, differing in only Ig2, only Ig3, or only Ig7. In these experiments, the pairs of variable domains were quite divergent, showing only 33%–50% amino acid sequence identity. In every case tested, homophilic binding was seen, and there was no sign of heterophilic binding. These experiments showed that all three of the variable domains contribute to binding, a result consistent with a separate experiment showing that the binding region resides within a fragment containing the first eight Ig-like repeats.

Not content to rest with these striking results, Wojtowicz et al. pushed the system even harder, testing Dscam isoforms that differed only in a single domain and where even that one domain was closely homologous, with 86%–92% sequence identity. Again, no prominent heterophilic binding was found. In only one case was weak heterophilic binding seen, for two closely related isoforms (7.27.25 and 7.27.26) that are 92% identical within the Ig7 domain. Even in this case, the assay of beads binding to cells was only weakly positive, and the coprecipitation assay remained negative, so even this one exception still supports the trend of a strong preference against heterophilic binding.

Without testing all 38,016 isoforms in all 1.4×10^9 pairwise combinations, it remains possible, of course, that some Dscam isoforms will show heterophilic binding. It is also possible that more sensitive quantitative assays might somewhat blur the line between the essentially all-or-nothing results from the assays performed so far. Nevertheless, this study seems to provide an impressively careful and unbiased screen of isoform interactions, and the results show a remarkably consistent and strong preference for homophilic binding specificity.

Functional Studies of Dscam in Drosophila

Neural Development

Earlier studies addressing the function of Dscam focused on loss-of-function alleles affecting all isoforms. The results demonstrated a broad requirement of Dscam throughout nervous system development, including the embryonic nervous system (Schmucker et al., 2000), the adult olfactory system (Hummel et al., 2003) as well as the mushroom body and ellipsoid body of the adult brain (Wang et al., 2002). These initial functional studies thus raised the exciting possibility that Dscam diversity could contribute to recognition not just in one region but throughout the nervous system. However, they did not

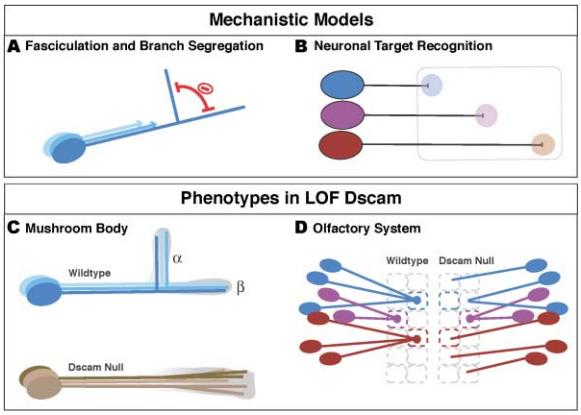

Figure 2. Models and Examples of the Functional Requirements of Dscam

(A) Repulsion of axonal sister branches might depend on the presence of identical Dscam isoforms. In contrast, expression of nonidentical isoforms might support adhesion of different axons in axon fascicles. (C) The sister branch repulsion model could explain collapse of mushroom body α/β lobes. (B) Given the binding specificity of individual isoforms, it is conceivable that matching up of specific combinations of isoforms on neurons and their targets may allow identification of precise targeting sites. (D) Such a model could be consistent with the functional requirement of Dscam in the olfactory system. In *Dscam* mutant animals, olfactory neurons often fail to connect to the correct target glomerulus (Hummel et al., 2003).

directly address the key question of whether isoform diversity is significant for Dscam function.

Two studies now published in *Neuron* examine in more detail the requirement of Dscam for mushroom body development and begin to explore the significance of diverse isoforms. The MB is the olfactory learning and memory system of insects and much is known about its stereotyped development. Experimentally, it is well suited for detailed clonal (i.e., mosaic) analysis, and axonal projections of single neurons can be visualized. Axonal branching and the formation of two bifurcating sister branches is a prominent feature of MB development and leads to the formation of distinct lobes (α and β lobes, see Figure 2C). Phenotypic abnormalities of axonal bifurcation can therefore be readily scored as malformations of the MB lobes. This was the basis for an earlier study showing that Dscam is important for sister branch segregation of MB neurons (Wang et al., 2002).

In a new study, Zhan et al. (2004) now show that Dscam is actually required at multiple steps during mushroom body development. High expression of Dscam can be found in young axons, which form a single distinct "core" bundle surrounded by a "bark" of older fibers. Functionally, complete loss of Dscam alleles results in multiple (i.e., split) core fibers, most likely arising from impaired fasciculation during early MB development. However, it should be noted that expressivity of this fasciculation phenotype is relatively low, suggesting that the role of Dscam may not be pivotal here and that other receptors can compensate for loss of Dscam.

More directly related to the question of diversity are experiments addressing which Dscam isoforms are expressed in MB neurons. Zhan et al. used previously described custom made oligo-arrays (Neves et al., 2004) to determine isoform expression in isolated MB neurons, sorted by using MB-specific GFP transgene markers, and found that the population of MB neurons express a large number of alternative exons. Despite the large number of exons expressed in MBs, the exon usage profile for one of the clusters, the exon 9 cluster, shows clear differences compared to expression in whole brain or photoreceptor neurons. Following this observation of restricted exon 9 expression, Zhan et al. determined whether single MB neurons express multiple exon 9 sequences. Indeed they found that for nine tested neurons each one expressed multiple exon 9 sequences. However, it should be pointed out that MB neurons are a heterogeneous population, so the expression profile of one neuron cannot be compared in these experiments to the exactly equivalent neuron in a sibling animal. It is therefore still unclear whether each neuron expresses a stochastically selected set of isoforms or whether specific neurons expresses a specific isoform subset that may correlate with their targeting specificity.

A series of subsequent experiments then address the question of whether it is important that MB neurons express a diverse set of Dscam isoforms. They took flies with null mutation of Dscam and performed rescue experiments of MB phenotypes using single cDNA transgenes. Two different Dscam isoforms each provided substantial rescue of MB defects, and no significant difference in rescuing ability was detected between them. These findings therefore do not confirm (but also do not exclude) a role for molecular diversity and may indicate a basic function of Dscam that does not depend on the specificity of individual isoforms.

In contrast to the ability to rescue defects by expressing single isoforms in a small number of neurons, the overexpression of a single isoform in many neurons created strong dominant phenotypes in MB neurons, similar to what has been shown earlier in embryonic photoreceptor neurons (Schmucker et al., 2000). The dominant phenotypes occurred equally well with transgenes expressing different extracellular domains. This may suggest either that it is important that different neurons express different Dscam isoforms or that they need to be present at the precisely correct concentration or developmental time.

Studies by Tzumin Lee's lab followed up on their previous report showing that segregation of sister branches of bifurcated MB neurons is disrupted in neurons that lack all Dscam. They approach the question of whether this function requires diverse Dscam isoforms by generating *Dscam* alleles that lack several (up to 9) of the 12 alternative exon 4 sequences. No detectable phenotypic abnormalities in segregation of sister branches of MB neurons were seen. It should be pointed out, however, that even in MB neurons lacking all Dscam function only some 35% of neurons show defects in sister branch formation. The sensitivity of detecting Dscam requirements in these neurons is therefore comparatively low and may not be well suited to detect subtle defects expected from loss of only a few alternative Dscam exons. It may be necessary to analyze more complex synaptic targeting processes, combined with a more substantial reduction of the alternative exon repertoire, to rigorously test the function of Dscam's molecular diversity.

In a series of experiments analogous to the studies of the Zipursky lab, Wang et al. attempted to rescue *Dscam* null embryos with minigene constructs containing only a single extracellular domain. The expression of this single isoform construct was driven by a fragment of the endogenous *Dscam* promoter controlling expression in the nervous system. This should be a technical improvement over standard GAL4/UAS expression systems. However, given the complexity of *Dscam*, it is at present not clear how accurately this promoter fragment reflects endogenous *Dscam* expression in respect to expression level or precise cellular distribution. Nevertheless, Wang et al. show that Dscam isoforms that contain the transmembrane domain 17.2, but not isoforms containing 17.1, are capable of rescuing the sister branch segregation of MB neurons. Both groups provide evidence that the transmembrane segment of Dscam may contain a signal that can influence subcellular localization to axons or dendrites.

In summary, the phenotypic analysis of mushroom body development examining loss and gain of *Dscam* function, as well as partial loss of diversity, makes it clear that at present no simple answer can be given about whether Dscam diversity is required for the correct specificity of neuronal wiring. On the one hand, forcing uniform expression of a single isoform has profound effects on neuronal wiring, suggesting a role for differential expression of diverse Dscam isoforms. On the other hand, substantial rescue can be achieved in some single neurons lacking all Dscam function by expressing only one isoform, questioning the specificity of individual isoforms. It appears clear that the complex receptor Dscam has multiple functions and that other experimental paradigms and requirements will have to be used to dissect the various functions allowing more definitive tests of the role of Dscam diversity.

Models and Conclusions

The recent studies of Dscam mark a breakthrough, showing how its extraordinary isoform diversity (Schmucker et al., 2000) may lead to a corresponding recognition diversity (Wojtowicz et al., 2004). Although a clear understanding of the in vivo role of diversity seems still out of reach, the homophilic isoform binding specificity is likely to provide important clues. One possibility is that matched combinations of pre- and postsynaptic receptor isoforms could specify neuron-target contacts (Figure 2B). This model seems compatible with results in the olfactory system, where *Dscam* loss-of-function disrupts the specificity of neuron-target interactions (Figure 2D; Hummel et al., 2003). A key unresolved question with this type of model is, however, that if a code of Dscam isoforms is indeed instructive, then the isoform expression presumably has to be prespecified in projecting neurons and target cells. So far, expression studies argue that single neurons contain many isoforms and that neurons of the same class express different and largely nonoverlapping sets, so this remains an open and pressing question.

Another model, not mutually exclusive with the one above, is that interaction of Dscam isoforms on adjacent axons may regulate their fasciculation (Figure 2A; Wang et al., 2002; Zhan et al., 2004). A particularly intriguing version of this model is that Dscam could mediate self-recognition, a property that may be crucial in neurons, as in the immune system, but which remains largely unexplored. In this case, Dscam might serve as a generator of random diversity, in which case the uniqueness of the isoform code is only required to distinguish self from nonself. In the MB, this could explain why *Dscam* loss of function results in failure of bifurcated axons to separate (Figure 2C). More generally, it might provide a way to help ensure—as is typically observed—that a single neuron's processes tend to spread out rather than crossing over or synapsing on itself.

Many other questions are raised by the current studies of Dscam. How does each *Dscam* mRNA select a single exon from the alternative arrays? How are homophilic Dscam interactions transformed into attractant or repellent responses? Could there be a special basis in Dscam structure or oligomerization that consistently favors homophilic binding despite massive isoform diversity? While much remains to be learned, the research has reached a turning point where generation of diversity in neural receptors has now moved from exciting models based on isoform diversity to the demonstration of a comparable molecular recognition diversity. It can be anticipated that future studies on Dscam will yield fascinating new insights into how a comparatively small gene number can give rise to neural connectivity in its immense complexity.

Selected Reading

Agarwala, K.L., Ganesh, S., Tsutsumi, Y., Suzuki, T., Amano, K., and Yamakawa, K. (2001). Biochem. Biophys. Res. Commun. *285*, 760–772.

Boulanger, L., Huh, G., and Shatz, C.J. (2001). Curr. Opin. Neurobiol. *11*, 568–578.

Dickson, B. (2002). Science *298*, 1959–1964.

Hummel, T., Vasconcelos, M.L., Clemens, J.C., Fishilevich, Y., Vosshall, L.B., and Zipursky, S.L. (2003). Neuron *37*, 221–231.

Missler, M., and Sudhof, T.C. (1998). Trends Genet. *14*, 20–26.

Neves, G., Zucker, J., Daly, M., and Chess, A. (2004). Nat. Genet. *36*, 240–246.

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). Cell *101*, 671–678.

Wang, J., Zugates, C.T., Liang, I.H., Lee, C.H., and Lee, T. (2002). Neuron *33*, 559–571.

Wang J, Ma, X., Yang, J.S., Zheng, X., Zugates, C, T., Lee, C.H., Lee, T. (2004). Neuron. 43(5), 663–672.

Wojtowicz, W.M., Flanagan, J.J., Millard, S.S., Zipursky, S.L., and Clemens, J.C. (2004). Cell *118*, 619–633.

Worby, C.A., Simonson-Leff, N., Clemens, J.C., Kruger, R.P., Muda, M., and Dixon, J.E. (2001). J. Biol. Chem. *276*, 41782–41789.

Wu, Q., and Maniatis, T. (1999). Cell *97*, 779–790.

Yamakawa, K., Huot, Y.K., Haendelt, M.A., Hubert, R., Chen, X.N., Lyons, G.E., and Korenberg, J.R. (1998). Hum. Mol. Genet. *7*, 227–237.

Zhan, X.L., Clemens, J.C., Neves, G., Hattori, D., Flanagan, J.J., Hummel, T., Vasconcelos, M.L., Chess, A., and Zipursky, S.L. (2004). Neuron *43*, 673–686.

Zhang, L.I. and Poo, M.M. (2001). Nat. Neurosci. Suppl. *4*, 1207–1214.